White Light Fringe Estimation: Algorithms, Error Sources and Mitigation Strategies

Mark Milman and Scott Basinger
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109
(818) 354-3627
Mark.H.Milman@jpl.nasa.gov and Scott.A.Basinger@jpl.nasa.gov

Abstract – This paper addresses the problem of highly accurate phase estimation at low light levels, as required by the Space Interferometry Mission (SIM). Most conventional phase estimation algorithms exhibit significant bias at the signal levels and requirements at which SIM will be operating. Several algorithms are analyzed, and methods for compensating for their bias are developed. Another source of error in phase estimation occurs because the phase is not constant over the integration period. Errors due to spacecraft motion, motion of compensating optical elements, and modulation errors are analyzed and simulated.

TABLE OF CONTENTS

1. INTRODUCTION

The Space Interferometry Mission (SIM) is a space-based long baseline optical interferometer designed to perform precision astrometry at unprecedented accuracy. A number of subsystems enable this performance. Among these is the starlight subsystem, which makes extremely accurate measurements of the phase difference of the starlight that enters the two arms of an interferometer. In the ideal setting, the combined light from the two arms of the interferometer is sent through a prism so that fringes formed at different wavelengths are dispersed linearly in wavenumber over a line of detector pixels. The relative optical path difference (OPD) between the two arms is varied by a piezo-electric actuator. Changes in intensity are measured at each pixel in the line detector. Therefore, the phase delay can be measured at many wavelengths, and the average of the all the phase delays divided by their wavenumbers gives an accurate estimate of the pathlength difference (the delay), the quantity of interest. A number of factors contribute to corrupting this procedure. This paper focuses on the analysis/development of the fundamental algorithms for white light fringe estimation at the low light levels experienced by the interferometer, and the effect and mitigation of several of these factors.

In order for SIM to make accurate astrometric measurements, the systematic errors in the computed delay measurements must be on the order of approximately 30 picometers. Delays are computed at approximately 1 ms time intervals, and are realized as the difference between an internal metrology measurement, which monitors the pathlength of the light through the optical system, and the white light fringe measurement. The relevant requirement on the instrument is that the average pathlength delay be computing to approximately this accuracy. One question addressed in the paper is whether the fringe measurement made over a 1 ms time interval is the true average phase (delay) over the period. The true delay is a time–varying quantity, and thus does not have a constant value even over the 1 ms period. In fact it can change by as much as several nanometers due to motions of the spacecraft and control elements. A related question is whether the underlying algorithms used for fringe position estimation are unbiased at these low light levels. And if so, what means are there for correcting them?

A careful assessment is made of both of these potentially deleterious phenomena. Various algorithms for estimating quasi-monochromatic light fringe positions, including the standard ABCD 4-bin algorithms and least squares algorithms using 4 and 8 bins are considered via simulation and analytical comparisons. The ABCD algorithm implemented in the simulations is essentially the four-bin method of dispersed white-light fringe detection that has been tested on the Palomar Testbed Interferometer [1,2], and which will also be implemented on the Keck Interferometer. It is shown that at low light levels, all of the algorithms discussed have a significant phase offset dependent bias, so that they cannot be implemented as is on SIM. Modifications to each of these algorithms to elim-

inate the bias are developed. These include both pointwise corrections for bias, and averaging methods that eliminate bias in time-averaged phase estimates. The efficacy and limitations of these modifications are validated in simulation studies. In addition, an analytical expression is derived for the pointwise estimation error incurred by a constantly varying phase. Large power low frequency disturbances are shown to produce a significant error from this effect. Strategies for minimizing these effects are discussed.

2. Monochromatic fringe estimation.

The instantaneous intensity at one wavelength is given as

$$I = I_0(1 + V\cos(kx + \phi)), \tag{1}$$

where I is the measured intensity, I_0 is the dc intensity term, V is the fringe visibility, k is the wavenumber of the monochromatic light, x is the dither length of the modulating element, and ϕ is the unknown phase to be estimated. In addition to ϕ, both V and I_0 are also unknown. Thus in general a minimum of three different values of x must be introduced to solve for these 3 unknown terms.

There are two ways of taking the intensity measurements in (1): (i) the phase stepping method where the phase shift is implemented and then the intensity is measured, and (ii) the integrating bucket method wherein measurements are made continuously as the phase is shifted [3]. Using either method, for the purposes of our initial studies, the intensity measurement can be assumed to have the form (1).

Suppose N ($N \geq 3$) steps are made so that the set of measurements

$$I_i = I_0(1 + V\cos(kx_i + \phi)) \tag{2}$$

is obtained. Introducing the function $F = [F_1, ..., F_N]$,

$$F_i(I_0, V, \phi) = I_0(1 + V\cos(kx_i + \phi)), \tag{3}$$

one straightforward way of estimating ϕ is to solve the nonlinear least squares problem

$$\min_{I,V,\phi} \sum |F_i(I_0, V, \phi) - I_i|^2. \tag{4}$$

A method similar to this is used in [5]; however, in [5] the additional parameters of wavenumber and dither stroke are also solved for in the least squares solution.

The nonlinear least squares problem posed in (4) can be circumvented by using the variables I_0, $I_0V\cos(\phi)$, and $I_0V\sin(\phi)$ as the unknown parameters [4]. The variables $I_0V\cos(\phi)$ and $I_0V\sin(\phi)$ are also referred to as *phasors.*

The motivation for this substitution comes from the observation that (2) can be written as

$$I_i = I_0(1 + V[\cos(kx_i)\cos(\phi) - \sin(kx_i)\sin(\phi)]), \tag{5}$$

so that the problem becomes linear in these variables. Specifically, (2) can be formulated as the linear system

$$\begin{bmatrix} I_1 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} 1 & \cos(kx_1) & -\sin(kx_1) \\ \vdots & \vdots & \vdots \\ 1 & \cos(kx_N) & -\sin(kx_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0V\cos(\phi) \\ I_0V\sin(\phi) \end{bmatrix}. \tag{6}$$

Now let $x = [I_0, I_0V\cos(\phi), I_0V\sin(\phi)]$, denote the vector of unknown variables. Then if $\hat{x}$ is any unbiased estimate of x, the estimate of the unknown phase $\hat{\phi}$ is obtained as

$$\hat{\phi} = \arctan\left(\frac{\hat{x}_3}{\hat{x}_2}\right). \tag{7}$$

(Here we have assumed that $-\pi/2 \leq \phi \leq \pi/2$.) Most phase estimation algorithms can be derived from these assumptions.

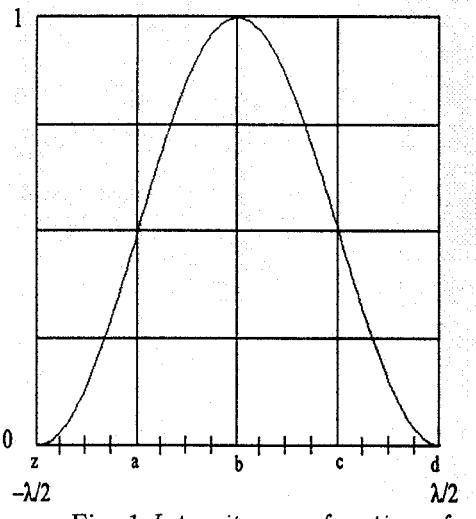

Fig. 1 *Intensity as a function of scan*

The nominal four-bin ABCD detection technique is one example of this approach. It works by scanning the OPD by exactly one full wave of the working wavelength, from -$\lambda/2$ to $\lambda/2$. The intensity is measured at five equal intervals during this scan. Here we are assuming that the PZT is actuated by a sawtooth waveform. Figure 1 shows the instantaneous intensity at the detector from a single wavelength during a continuous scan from -$\lambda/2$ to $\lambda/2$. The values measured for the four-bin method are

the integrals of this plot at five equally-spaced intervals, shown as z, a, b, c, and d on Fig 1.

The integrated flux in each bin is calculated as $A = a - z$, $B = b - a$, $C = c - b$ and $D = d - c$. Introducing the variables $X = A - C$ and $Y = B - D$, the phase delay is calculated as

$$\hat{\phi} = \arctan\left(\frac{X+Y}{Y-X}\right), \tag{8}$$

Another form for the calculation is

$$\hat{\phi} = \arctan\left(\frac{B-D}{A-C}\right) - \frac{\pi}{4}. \tag{9}$$

Modifications to (8) and (9) are necessary when the stroke length and wavelength are unequal [1]. For example, (8) is modified as

$$\hat{\phi} = \arctan\left(\gamma\frac{X+Y}{Y-X}\right), \tag{10}$$

where

$$\gamma = \frac{2\sin(s/4) - \sin(s/2)}{1 - \cos(s/2)} \tag{11}$$

and s is the product of the stroke length and wavenumber.

Other unbiased estimators for determining the phasor and intensity parameters include least squares and minimum variance estimators. The distinction between the two is that the minimum variance estimate incorporates the statistics of the measurement errors of I_i. We note that the least squares and minimum variance solutions extend to an arbitrary number of dither steps which are not required to be of equal length. There is also a simple relationship between the solution to the linear and nonlinear least squares problems ((6)–(7) and (4)). Let Ψ denote the diffeomorphism mapping the variables $\{I_0, I_0V\cos(\phi), I_0V\sin(\phi)\}$ into $\{I_0, V, \phi\}$. Then the composite function $F \circ \Psi$ is the matrix in (6). Hence, the solutions to the two problems are the same.

3. Analysis of Methods

This section focuses on developing approximate bias and variance estimates for the algorithms discussed above. Bias corrections are developed, and simulations are performed to assess the efficacy of the algorithms.

Estimator Bias

The first question we address is whether these estimators are biased. This question is very important from the standpoint of the astrometric objectives of fringe estimation since the quantity of interest is the *average* phase over the period of an observation of a science object. The typical length of such an observation is about 30sec, during which time as many as 30,000 phase calculations are made. The average error in these calculations is required to be on the order of 30 picometers. This requirement translates to the technical question "does $E(\hat{\phi}) = \phi$ where $\hat{\phi}$ is given by (7), (8), or (9)?" This question is taken up now.

Without loss of generality we may assume that once an unbiased estimator has been chosen for the variables $I_0, I_0V\cos(\phi), I_0V\sin(\phi)$, the phase estimate $\hat{\phi}$ has the general form

$$\hat{\phi} = \arctan\left(\frac{\sum h_{1j}I_j}{\sum h_{2j}I_j}\right), \tag{12}$$

where h_{ij} are the gains for the linear estimator. Each I_i is a random variable that is modeled as a sum of 2 independent random variables; one with a Poisson distribution corresponding to shot noise, and the other with a Gaussian distribution corresponding to read noise.

Because the linear estimates are unbiased, it is true that

$$\phi = \arctan\left(\frac{\sum E(h_{1j}I_j)}{\sum E(h_{2j}I_j)}\right), \tag{13}$$

but this does not imply $E(\hat{\phi}) = \phi$. To compute $E(\hat{\phi})$, let $p_i(x_i)$ denote the density function of the random variable I_i, and let $\bar{x}_i$ and $\sigma^2_{x_i}$, denote its mean and variance, respectively. Then since the I_i's are independent,

$$E(\hat{\phi}) = \int\cdots\int \arctan\left(\frac{\sum h_{1j}x_j}{\sum h_{2j}x_j}\right) p_1(x_1)\cdots p_N(x_N)dx_1\cdots dx_N. \tag{14}$$

Now expand the arctan function to second order about the means $\bar{x}_1, ..., \bar{x}_N$ to obtain

$$E(\hat{\phi}) = \phi + \frac{1}{2}\left\{\sum \frac{\partial^2}{\partial x_i^2}\arctan\left[\frac{\sum h_{1j}x_j}{\sum h_{2j}x_j}\right]\sigma^2_{x_i}\right\}, \tag{15}$$

where all of the derivatives are evaluated at the mean values $\bar{x}_i$.

The approximation to the bias in the estimate is contained in the second term on the right in (15). For example, these considerations can be used to generate a bias correction term to the nominal ABCD algorithm when there is a mismatch between the wavelength and stroke length. This correction can be computed as

$$\begin{aligned} b &= (2\sigma^2_{read} + A + C)\frac{2\gamma(1-\gamma^2)Y^3 - 4\gamma XY(2Y-(1+\gamma^2)X)}{\left[(Y-X)^2+\gamma^2(X+Y)^2\right]^2(Y-X)} \\ &+ 2(\sigma^2_{read} + B + D)\frac{2\gamma(1-\gamma^2)X^3 - 4\gamma XY(2X-(1+\gamma^2)Y)}{\left[(Y-X)^2+\gamma^2(X+Y)^2\right]^2(Y-X)} \end{aligned} \tag{16}$$

The figure below contains histograms of Monte Carlo simulations of estimating a single phase using the nominal ABCD algorithm and the bias correcting algorithm. Phase estimates are computed every millisec, and the simulation is run for either 1sec or 100sec. The simulations are then repeated 1000 times to produce the histograms. The wavelength of the source is 900nm, while the length of the PZT stroke is 725nm. It is seen that without bias correction the nominal ABCD algorithm exhibits a bias error of approximately .3nm. This error is about .014nm for the corrected algorithm.

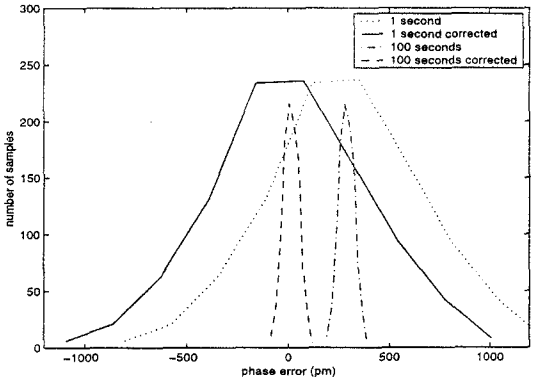

Fig. 2 *Bias in phase error estimated due to shot noise*

The ABCD algorithm does not need correction when the wavelength and stroke length match, but when using dispersed fringe measurements where the light is spread out over several spectral bins, this condition cannot be satisfied. (Although see [1] for how this was handled on the Palomar Interferometer.)

We note that after computing the partial derivatives in (15) and writing $\sigma_{x_i}^2$ as the sum of read noise and shot noise components, it can be shown that the bias due to read noise decreases *quadratically* with the number of photons, while the bias due to shot noise decreases *linearly* with the number of detected photons. Thus, the magnitude of the bias decreases with the strength of the signal, and ultimately becomes inconsequential at high enough light levels. Note also that the bias is in general a function of the phase offset.

A somewhat more direct approach to developing an unbiased estimate of the phase is to approximate the conditional mean $\hat{\phi}_{opt} = E(\phi|I_1,...,I_N)$. Recall that the conditional mean is the unbiased minimum variance estimate of the phase offset ϕ given the observations $\{I_1,...,I_N\}$. ϕ may be written as a function of the variable $x = [I_0, I_0V\cos(\phi), I_0V\sin(\phi)]$ as

$$\phi(x) = \arctan\frac{x_3}{x_2}. \tag{17}$$

Thus,

$$E(\phi|I_1,...,I_N) = \int \phi(x_1,x_2,x_3)p(x_1,x_2,x_3|I)dx_1dx_2dx_3, \tag{18}$$

where $p(x|I)$ is the conditional density of x given I. We now expand ϕ in a Taylor series about the conditional mean of x given I to obtain

$$\begin{aligned}\hat{\phi}_{opt} &\approx \arctan\left(\frac{\hat{x}_3}{\hat{x}_2}\right) + [p_{22}-p_{33}]\frac{\hat{x}_2\hat{x}_3}{(\hat{x}_2^2+\hat{x}_3^2)^2} \\ &+ p_{23}\frac{\hat{x}_2^2-\hat{x}_3^2}{(\hat{x}_2^2+\hat{x}_3^2)^2}\end{aligned} \tag{19}$$

Where p_{ij} is the covariance matrix of the minimum variance solution in (6). In this approach the bias correction terms (the latter two terms on the right above) appear very naturally, and are easily computed using the covariance matrix obtained from the linear estimates of the dc intensity and phasor variables, I_0, $I_0V\cos(\phi)$, and $I_0V\sin(\phi)$, respectively.

There is still another approach that may be used to eliminate bias. This is based on the recognition that the 30pm requirement applies to the *averaged* phase estimate. Therefore it is actually not required to remove the bias from each phase estimate; just as long as it is removed from the averaged estimate. An approach for doing this is to first average the phasors, $I_0V\cos(\phi), I_0V\sin(\phi)$, and then take the arctan of the quotient [6]. The viability of this approach stems from the following observation:

Let $\{\phi_i\} \subset [-\pi/2, \pi/2]$ and define

$$\bar{\phi} = \frac{1}{N}\sum_{i=1}^{N}\phi_i, \overline{\sin(\phi)} = \frac{1}{N}\sum_{i=1}^{N}\sin(\phi_i), \overline{\cos(\phi)} = \frac{1}{N}\sum_{i=1}^{N}\cos(\phi_i) \tag{20}$$

Then

$$\bar{\phi} = \arctan\left\{\frac{\overline{\sin(\phi)}}{\overline{\cos(\phi)}}\right\} + \frac{1}{6N}\cos(2\bar{\phi})\sum_{i=1}^{N}\delta\phi_i^3 + O(|\delta\phi_i|_\infty^4), \tag{21}$$

where $\delta\phi_i = \bar{\phi} - \phi_i$.

This result can be interpreted in the following way: If $\{X_i\}_{i=1,N}$ and $\{Y_i\}_{i=1,N}$ are random variables with

$$E(X_i) = \kappa\sin(\phi_i), \quad E(Y_i) = \kappa\cos(\phi_i), \tag{22}$$

for some constant κ, then (retaining terms through third order in $\delta\phi_i$),

$$\bar{\phi} = \arctan\left(\frac{E\sum_{i=1}^{N}X_i}{E\sum_{i=1}^{N}Y_i}\right) + \frac{\cos(2\bar{\phi})}{6N}\sum\delta\phi_i^3. \tag{23}$$

Note that if the $\delta\phi_i$ are random samples from a symmetric distribution, then the contribution of the third order error can also be ignored for large values of N.

This approximation allows us to use the average values of the phasors for estimating the average phase $\bar{\phi}$ for the following model. Suppose the "true" phase is a step function that takes on the discrete values ϕ_i on the time interval $[t_i, t_{i+1}]$. Further assume that the dc intensity values and visibilities are independent of i. Then beginning with *any* unbiased algorithm for estimating phasors, an asymptotically unbiased algorithm (through the second order variations in $\delta\phi_i$) for estimating $\bar{\phi}$ is obtained by simply averaging the phasors and then computing the arctan of the resulting ratio.

Estimator Variance

Assuming there is no bias in the estimate, the variance of the estimate is an important consideration for astrometric performance because it dictates how much integration time is required to achieve a certain level of precision in the average phase estimate.

To compute the variance of the least squares or minimum variance estimate we perform calculations similar to those in (19–20). Let ϕ denote the true phase and let $\hat{\phi}$ be the estimate. Then,

$$E(|\phi - \hat{\phi}|^2) = \frac{p_{xx}\cos^2(\phi) + p_{yy}\sin^2(\phi) - p_{xy}\sin(2\phi)}{I_0^2 V^2}. \tag{24}$$

Here p_{xx}, p_{yy}, p_{xy} form the covariance matrix of the phasor variables from any linear estimator for (6). (In particular these expression apply to both the optimal and least squares estimates.)

The variance for the ABCD algorithm is given as

$$E(|\hat{\phi}-\phi|^2) = \int\int [\arctan\left(\gamma\frac{x+y}{y-x}\right) - \phi]^2 p_X(x) p_Y(y) dxdy, \tag{25}$$

where the factor γ that accommodates the mismatch is given in (11). Now let $f(x,y) = \arctan(\gamma(x+y)/(y-x))$. Then using the identity

$$\phi = \arctan\left(\gamma\frac{\bar{x}+\bar{y}}{\bar{y}-\bar{x}}\right), \tag{26}$$

where $\bar{x} = E(X)$ and $\bar{y} = E(Y)$, we approximate the integrand in (24) by a Taylor series expansion of f about $(\bar{x}, \bar{y})$ and retain terms containing the second moment to obtain the approximate variance expression:

$$E(|\hat{\phi} - \phi|^2) = \left[\frac{\partial f}{\partial x}\right]^2 \sigma_X^2 + \left[\frac{\partial f}{\partial y}\right]^2 \sigma_Y^2. \tag{27}$$

In the figure below analytical comparisons are made between these algorithms with respect to rms phase error due to shot noise (240 photons per observation). The wavelength and PZT stroke length are both 725nm.

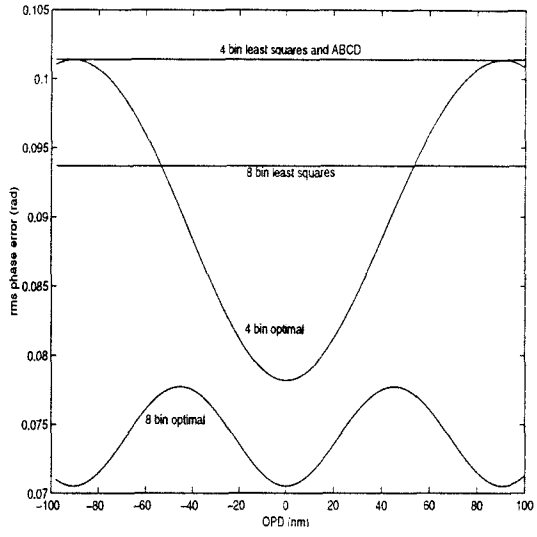

Fig. 3 *RMS error of estimators*

It is seen that the "optimal" algorithms perform better than the least squares or ABCD algorithms.

Simulation results

The results of a number of simulations with the ABCD and optimal 4–bin and 8–bin algorithms are tabularized below. The simulations are developed in an integrating bucket fashion [3] with a stroke length of 900nm. The bias and variance of each of the algorithms is investigated with respect to varying wavelength of the light. Wavelengths of 550nm, 625nm, 725nm, and 900nm are used. The photon counts are generated by sub-sampling the OPD dithering steps to calculate the integral which determines the expected value of the number of photons for each bin. Shot noise and read noise are then independently added. The shot noise is governed by a Poisson distribution with a mean of 240 photons, and the read noise is zero mean Gaussian with a one sigma value of 6 photons. A visibility of one is used. The statistics for each of the results is calculated by computing the mean and standard deviation of 1000 simulations, with each simulation consisting of 1000 phase estimates representing the number of phase estimates the interferometer will produce in one second.

The results indicate that the optimal algorithms perform significantly better than the ABCD or phasor averaging methods when there is a larger mismatch between the operating wavelength of the light and the stroke length. In

fact a factor of greater than 5 improvement in standard deviation is seen at the shortest wavelength. When the wavelength and stroke length are matched the optimal 8 bin algorithm performs slightly worse than all the algorithms because of the additional read noise introduced by the extra four bin measurements it makes to generate an estimate.

Global parameters:
nsim = 1000; % number of Monte Carlo runs
npoi = 1000; % length of individual data run (ms)
rn = 8; % read noise
photons =240; % total number of photons in measurement
wavelen=900; % wavelength piezo movement was optimized for (nm)

visWL = 1; % visibility
phase_delay = 0.01*pi; % phase delay (radians)
offwave=900; % actual wavelength data was taken at (nm)

	Shot Error		Read Error		Total Error	
	Mean	Std	Mean	Std	Mean	Std
ABCD, bias estimator	-9.64	465.51	-12.42	359.24	-19.8	592.71
average phasors	-9.85	453.38	-13.46	358	-23.06	584.5
optimal 4-bin	-15.29	364.12	-12.42	359.24	-25.43	553.36
optimal 8-bin	7.04	324.54	-16.49	466.82	-18.7	606.06

visWL = 1; % visibility
phase_delay = 0.01*pi; % phase delay (radians)
offwave=725; % actual wavelength data was taken at (nm)

	Shot Error		Read Error		Total Error	
	Mean	Std	Mean	Std	Mean	Std
ABCD, bias estimator	16.32	440.38	4.57	342.87	22.38	582.87
average phasors	16.75	442.03	4.49	342.2	21.25	585.4
optimal 4-bin	10.77	415.33	7.72	296.17	22.31	578.44
optimal 8-bin	-10.85	236.71	9.45	342.17	-7.24	472.75

visWL = 1; % visibility
phase_delay = 0.01*pi; % phase delay (radians)
offwave=550; % actual wavelength data was taken at (nm)

	Shot Error		Read Error		Total Error	
	Mean	Std	Mean	Std	Mean	Std
ABCD, bias estimator	167.72	1286.36	98.22	957.24	302.25	1591.83
average phasors	-47.74	1311.92	14.04	969.66	-33.92	1637.27
optimal 4-bin	3.42	240.29	12.1	199.44	13.9	327.78

4. Errors due to change in OPD.

The previous section focused on statistical properties of various estimators, given that the phase that is being estimated is constant over the period of integration. Because the SIM requirements on the estimation of the pathlength are subnanometer, it is not accurate to make this assumption. Factors that contribute to a non–constant phase include instrument motion (both rigid body and vibrational motions), errors introduced by the dither signal, and errors introduced by the control system while trying to reduce the phase errors.

To begin we make a slight modification of the model (2), explicitly using a bucket integrating method where photons are collected contemporaneously during the dither of the modulating element. And we also allow ϕ to be time-varying during the integration period. The resulting intensity model is

$$I_i = \int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}} I_0\{1+V\cos(u+\phi(u))\}du, \qquad (28)$$

where $u = kx$, u_i is the central value for each bucket, and the buckets have equal widths Δ. If ϕ is constant, the model becomes after integrating (6)

$$I_i = I_0\{\Delta + 2V\sin(\frac{\Delta}{2})\cos(u_i+\phi)\}. \qquad (29)$$

Using the variables $I_0, I_0\tilde{V}cos(\phi), I_0\tilde{V}\sin(\phi)$ where $\tilde{V} = 2\sin(\Delta/2)V$, the system of equations above can be written as

$$\begin{bmatrix} I_1 \\ \vdots \\ I_N \end{bmatrix} = \begin{bmatrix} \Delta & \cos(u_1) & -\sin(u_1) \\ \vdots & \vdots & \vdots \\ \Delta & \cos(u_N) & -\sin(u_N) \end{bmatrix} \begin{bmatrix} I_0 \\ I_0\tilde{V}\cos(\phi) \\ I_0\tilde{V}\sin(\phi) \end{bmatrix}, \qquad (30)$$

just as in (6). For notational purposes we write this system as

$$I = Ax. \qquad (31)$$

When ϕ is not constant, let $\bar{\phi}$ denote its mean value and define the variation $\delta\phi$ by $\phi = \bar{\phi} + \delta\phi$. Note that

$$\sum \int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}} \delta\phi(u)du = 0. \qquad (32)$$

In the analysis to follow we assume that terms of $O(|\delta\phi^2|)$ and higher can be ignored. Thus we have

$$\begin{aligned} I_i &= \int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}} I_0\{1+V\cos(u+\bar{\phi}+\delta\phi(u))\}du \\ &= I_0\{\Delta + 2V\sin(\tfrac{\Delta}{2})\cos(u_i+\bar{\phi})\} \\ &\quad -I_0V\cos(\bar{\phi})\int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}}\sin(u)\delta\phi(u)du \\ &\quad -I_0V\sin(\bar{\phi})\int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}}\cos(u)\delta\phi(u)du\} \end{aligned} \qquad (33)$$

Next introduce the matrix B, with entries $B_{ij} = 0$ if j=1,

$$B_{ij} = \frac{1}{2\sin(\frac{\Delta}{2})}\int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}}\sin(u)\delta\phi(u)du \qquad (34)$$

if j=2 and

$$B_{ij} = \frac{1}{2\sin(\frac{\Delta}{2})}\int_{u_i-\frac{\Delta}{2}}^{u_i+\frac{\Delta}{2}}\cos(u)\delta\phi(u)du \qquad (35)$$

if j=3.

Then (34) has the form

$$I = Ax - Bx. \qquad (36)$$

Now suppose an estimate $\hat{x}$ of x is generated via

$$\hat{x} = (A^TQ^{-1}A)^{-1}A^TQ^{-1}I. \qquad (37)$$

Then since

$$x = (A^T Q^{-1} A)^{-1} A^T Q^{-1} (I + Bx), \tag{38}$$

the error $e = x - \hat{x}$ is simply

$$e = (A^T Q^{-1} A)^{-1} A^T Q^{-1} Bx. \tag{39}$$

And because the phase is given by

$$\phi(x) = \tan^{-1}(x_3/x_2), \tag{40}$$

the phase error to first order is

$$\begin{aligned} \phi(x) - \phi(\hat{x}) &= \phi'(x)(x - \hat{x}) \\ &= \phi'(x)e \\ &= \frac{-x_3}{x_2^2+x_3^2} e_2 + \frac{x^2}{x_2^2+x_3^2} e_3 \\ &= \frac{-\sin(\bar{\phi})e_2}{I_0 \tilde{V}} + \frac{\cos(\bar{\phi})e_3}{I_0 \tilde{V}}. \end{aligned} \tag{41}$$

Equation (41) gives the general form of the phase error. To obtain a more useful characterization it is necessary to evaluate the error vector e given in (39). Two simplifying assumptions are made to facilitate this analysis. These are: (1) Q is a scalar matrix (a multiple of the identity), and (2) the wavelength and dither stroke are matched. Under these assumptions an exact formula for the error can be derived when $\delta\phi$ is a step function with values $\delta\phi_i$ for $u \in [u_i - \Delta/2, u_i + \Delta/2]$. Specifically, we find that

$$e_2 = \frac{I_0 \tilde{V}}{N} [\cos(\bar{\phi}) \sum \delta\phi_i \sin(2u_i) + \sin(\bar{\phi}) \sum \delta\phi_i \cos(2u_i)], \tag{42}$$

and

$$e_3 = \frac{I_0 \tilde{V}}{N} [\cos(\bar{\phi}) \sum \delta\phi_i \cos(2u_i) - \sin(\bar{\phi}) \sum \delta\phi_i \sin(2u_i)], \tag{43}$$

where N is the number of dither steps. The resulting error in the phase estimate is

$$\begin{aligned} \phi(x) - \phi(\hat{x}) &= \tfrac{1}{N} [\cos(2\bar{\phi}) \textstyle\sum \delta\phi_i \cos(2u_i)] \\ &\quad - \sin(2\bar{\phi}) \textstyle\sum \delta\phi_i \sin(2u_i)]. \end{aligned} \tag{44}$$

Assuming $\bar{\phi} < \pi/8$ (which is a reasonable assumption while fringe tracking is taking place), the error is maximized when

$$\delta\phi_i = |\delta\phi|_\infty \cos(2u_i), \quad \text{where} \quad |\delta\phi|_\infty = \max_i |\delta\phi_i|, \tag{45}$$

with resulting error $|\delta\phi|_\infty/2$.

If instead the $\delta\phi_i$ are treated as identically distributed independent random variables with zero mean and variance σ^2, we have

$$\phi(x) - \phi(\hat{x}) = \frac{1}{N} [\cos(2\bar{\phi} C_2 \quad - \sin(2\bar{\phi}) S_2] \delta\phi, \tag{46}$$

where

$$C_2 = [\cos(2u_1), \cos(2u_2), \ldots, \cos(2u_N), \tag{47}$$

and

$$S_2 = [\sin(2u_1), \ldots, \sin(2u_N)]. \tag{48}$$

Thus,

$$\begin{aligned} E(\mid \phi(x) - \phi(\hat{x}) \mid^2 &= \frac{\sigma^2}{N^2} tr\{C_2 C_2^T + S_2 S_2^T\} \\ &= \frac{\sigma^2}{N^2}. \end{aligned} \tag{49}$$

Note that when we assume that $\delta\phi_i$ is random, increasing the number of dither steps leads to a linear decrease in the rms error. But this is not the case in general, as seen in (45). These results hold regardless of the mechanism that produces the phase change over the integration period; and thus can also be interpreted as error in the dither position. From this perspective our results confrom with those obtained in [7].

Equation (44) also has the following approximate continuous analogue when $\delta\phi$ is not a step function:

$$\begin{aligned} \phi(x) - \phi(\hat{x}) \quad \approx \quad & \tfrac{1}{2N \sin(\frac{\Delta}{2})} \{ - \sin(2\bar{\phi}) \textstyle\int_0^{2\pi} \sin(2u) \delta\phi(u) du) \\ & + \cos(2\bar{\phi}) \textstyle\int_0^{2\pi} \cos(2u) \delta\phi(u) du \}. \end{aligned} \tag{50}$$

Thus it is seen that the error introduced by the nonconstant phase term is completely characterized by its 2nd Fourier component.

Another useful error characterization can be developed if we suppose that over each integration period the deviation, $\delta\phi$, from the mean phase is approximated by a quadratic. Without loss of generality we will assume that $u \in [-\pi, \pi]$ in the analysis. Then

$$\delta\phi(u) = \delta\phi(0) + \delta\phi'(0)u + \frac{1}{2}\delta\phi''(0)u^2. \tag{51}$$

It can be shown that for functions of quadratic type the approximation (44) is quite accurate. A straightforward integration gives the error as

$$\phi(x) - \phi(\hat{x}) \approx \frac{\pi}{2N \sin(\Delta/2)} [\cos(2\bar{\phi}) \frac{\delta\phi''(0)}{2} - \sin(2\bar{\phi}) \delta\phi'(0)]. \tag{52}$$

Two examples of expected phase motion produced by the instrument that illustrate the results of this section are now considered. The first involves OPD variations induced by the rigid body motion of the spacecraft, and the second involves the phase variations produced by the control delay line PZT motion while compensating for pathlength error.

A simulation was first developed to determine the effect of undetected rigid body motion of the SIM spacecraft on the fringe measurement process. The motion of the spacecraft is assumed to be of very low frequency compared to the 1kHz measurements of the delay line fringes. Therefore the change in phase during a 1ms time interval is approximated by a ramp function. Using nominal SIM attitude control system values of 2arcsec pointing stability with a .1Hz controller bandwidth, the maximum OPD variation during a millisec interval can be as large as 10nm. The simulations used for these plots contained no noise, and the ABCD algorithm with matched stroke and wavelength of 725nm was employed. The form of the controller assumed in these simulations also contributes in an important way to the observed error. In these simulations the PZT is moved instantaneously to the commanded compensating position. (Some comments on this form of the controller using a zero order hold assumption will be made after discussing the results of the rigid body simulations.)

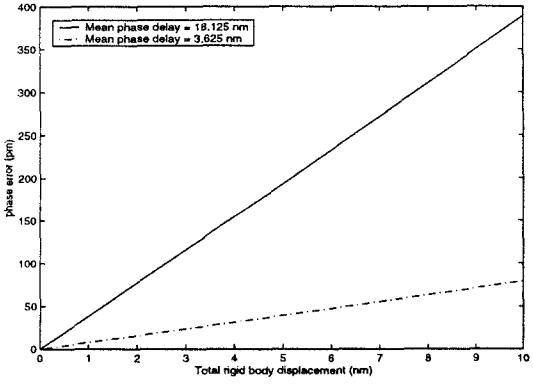

Fig. 4 *Systematic phase estimation error due to spacecraft motion*

The abcissa in the plot can be interpreted as the OPD velocity: The 10nm value can be related as a .1Hz bandwidth ACS controller, while a 1nm value corresponds to a .01Hz bandwidth controller. Note that the error in phase measurement is dependent on how well phased the two arms of the inteferometer are to begin with. The two separate plots show two different starting phases. If the starting phase is 0nm, there is no error in the measurement, but we cannot rely on this being the case. The larger the initial phase, the more error the rigid body motion produces. For each of the two initial (or mean) phases, several data points were collected for different overall motions of the spacecraft. The plots show that the phase error is linear with the overall motion. These results conform very well to the prediction made in (51). Because the acceleration term is zero for the ramp function, (51) states that the error should be proportional to the product of the phase velocity and the initial phase error; exactly what is observed in the simulations.

These results indicate that both the rigid body motion of the spacecraft should be minimized as well as the initial phase offset between the two arms of the interferometer to avoid systematic errors in the phase measurements. However, low frequency error such as this can also be mitigated by modifying the implementation of the controller. This involves running a higher bandwidth inner-loop that moves the delay line PZT to track the OPD rate. For example, a 5Khz inner-loop run in this manner should reduce the error in the plots by nearly a factor of 5.

The next set of plots illustrate another source of phase measurement error. As in the previous set, we assume the spacecraft rotates, causing a 10nm change in OPD over 1ms. The interferometer measures the change in OPD and the delay line control system applies a correction once per per millisec to compensate for it. The finite response time of the actuator due to electronics will have an effect on the estimated phase. This response is modeled in the following plots via a damping coefficient a:

$$u(t) = u_0(1 - \exp(-at)), \tag{53}$$

where u_0 is the desired correction. Figure xx shows the changing phase due to a 10nm correction being applied to the delay line at the beginning of the 1ms interval. Three different profiles, each with a different damping coefficient, are shown.

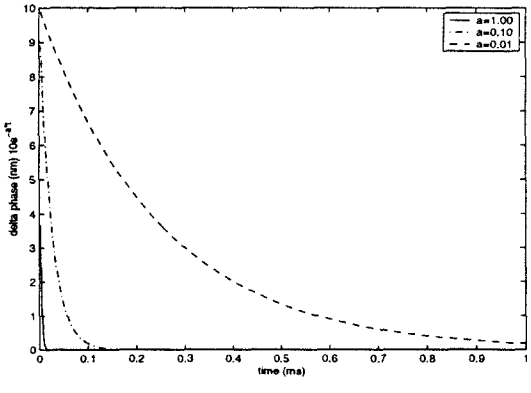

Fig. 5 *Phase profiles due to different response models*

Several different damping coefficients were used to generate the plots in Figure 6.

In each case the mean phase delay was 18.125nm and the total correction applied was 10nm. The examples show problems specific to SIM and therefore have typical parameters that we expect to encounter. The total number of photons collected in each 1ms interval is 240. The

read noise in each bin is equivalent to 9 photons. The visibility was taken as unity, and the wavelength of the light used is 725nm. An average of 10^6 measurements are used to determine the errors shown. For large damping coefficients (corresponding to fast electronics), the error due to finite response times is minimal, less than 10pm. However, the errors increase as the electronics get slower.

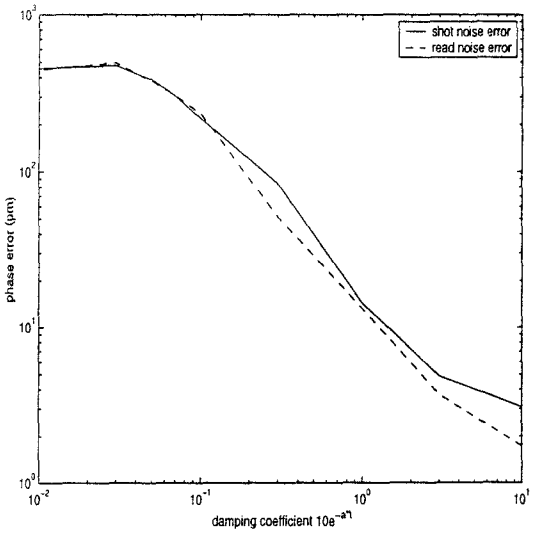

Fig. 6 *Error in estimated phase due to actuator response*

5. Concluding Remarks

The Space Interferometry Mission poses interesting challenges in phase estimation due to the required precision and low light levels at which the instrument operates. In this paper we showed that conventional phase estimation techniques do not in general meet SIM's requirements due to small inherent bias in each of these methods at low signal signal levels A general technique for reducing the bias was developed, and the efficacy of this modification was validated in simulations. Comparisons of the ABCD, least squares and minimum variance based phase estimation algorithms were conducted using the anticipated operating parameters of the instrument. The relative performance of these estimators was shown to be a function of wavelength; but the 8 bin minimum variance based algorithm was shown to be generally superior to the others. We also analyzed and simulated the effect of non-constant pathlength difference on phase estimation. This was done in the context of pathlength difference variations due to spacecraft motion and in the motion of the PZT element used to control the pathlength. Each of these effects was shown to impact the phase estimation problem, but in predictable ways that can be compensated.

The analysis and simulations presented were restricted to quasi–monochromatic light. Current and future work focuses on extending the work to multiple spectral bins of finite bandwidth. In addition, the simulations are being extended to include the closed loop delay line controller.

Acknowledgements

This work was prepared at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

[1] Colavita, M. M., et al.,"The Palomar Testbed Interferometer," The Astrophysical Journal 510, 505-520 (1999).

[2] Colavita, M. M., "Fringe visibility estimators for the Palomar Testbed Interferometer," PASP 111, 111-117 (1999).

[3] K. Creath, "Phase–Measurement interferometry techniques," in *Progress in Optics* XXVI, Elsever Science Publishers B. V., 1988.

[4] J. E. Grievenkamp, "Generalized data reduction for heterodyne interferometry," Opt. Eng., Vol. 23, July 1984, pp. 350–352.

[5] Kuhnert, A. C., Shaklan, S. B., Shen, T. P., "First Tests of the Interferometer in the the Micro-Arcsecond Metrology Testbed (MAM)," SPIE (Munich 2000).

[6] A. Quirrenbach, D. Mozurkewich, D. F. Buscher, C. A. Hummel, and J. T. Armstrong, "Phase-referenced visibility averaging in optical long-baseline interferometery,"in SPIE Miletone Series, Volume MS 139, Ed. Peter R. Lawson, Washington, 1997 (pp 578-586).

[7] J. Schwider, R. Burrow, K.–E. Elsner, J. Grzana, R. Spolaczyk, and K. Merkel, "Digital wavefront measuring interferometry: Some systematic error sources," Applied Optics, Vol 22, Nov. 1983, pp. 3421–3432.

Biographies

Mark Milman received the Ph.D. in mathematics from the University of Southern California in 1980. He has been at JPL since 1980, although he returned to USC as a visiting professor in the Department of Mathematics in the 1989-1990 academic year. He currently works for the Space Interferometry Mission project in which he performs various instrument system analyses, especially as it pertains to astrometric performance.

Scott Basinger received his Ph.D. in Electrical Engineering from the University of Illinois at Champaign-Urbana in 1996. Since then he has been working at the Jet Propulsion Laboratory in Pasadena, California, where he is a senior member of the technical staff and works on optical modeling, simulations, and wavefront sensing and control for adaptive optics.